

Installazione LSF

Rosalba Galloni

CILEA, Segrate

Abstract

Il 12 Maggio 1999 è stato installato al Cilea, sulle macchine HP Exemplar 2000 (icl spp2.cilea.it) e K580 (icl580.cilea.it) il sistema batch LSF (Load Sharing Facility) distribuito dalla Platform Computing Corporation (www.platform.com).

Al Cilea è stata installata la versione LSF Standard Edition la quale comprende i moduli Base e Batch. Questo prodotto ha sostituito il sistema batch NQS precedentemente attivo sull' Exemplar 2000.

Il sistema batch LSF installato al Cilea è un sistema di code che permette di integrare e distribuire il carico su più macchine.

Attualmente le macchine coinvolte sono le macchine HP Exemplar 2000 e K580.

Una delle novità del prodotto è la possibilità, automatica, di prevedere una soglia di carico ottimale della macchina (chiamato loadsched), al di sopra del quale non viene schedulato più nessun job.

E' possibile poi prevedere una seconda soglia di carico (loadstop) raggiunta la quale non solo non viene schedulato più nessun job, ma job running vengono sospesi in attesa che il carico si abbassi.

Nel primo caso il job risulterà in stato di PEND, nel secondo caso in stato di SUSP.

Tutto ciò è comunque trasparente per l'utente poiché viene fatto automaticamente dal sistema LSF.

Questo sistema batch è stato implementato per bilanciare e distribuire meglio il carico della macchina Exemplar 2000, che ultimamente è molto aumentato come si può anche notare dal grafico di utilizzo di tale macchina nel mese di aprile 1999.

Per rendere il più facile e meno traumatico all'utente l'utilizzo del nuovo sistema batch si è deciso di non cambiare i nomi delle code batch precedenti, ad esse si sono solo aggiunte le nuove code per la macchina K580.

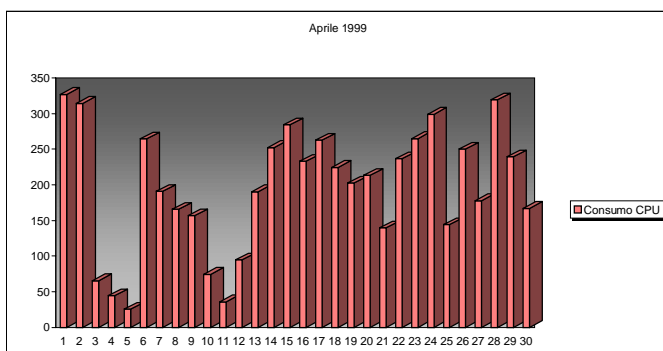
Nella tabella 1 si elencano il nome della coda, la macchina su cui vengono eseguiti i job, i limiti di tempo definiti sulla coda ed il subcomplex a cui punta nel caso la macchina sia l'Exemplar 2000.

Tabella 1

Nome Coda	Macchina	Limiti CPU	Subcomplex
short1	icl spp2	15 min.	sub1
short2	icl spp2	15 min.	sub2
short_k580	icl580	15 min.	-----
batch_seq	icl spp2	None	sub1
inf	icl spp2	None	sub1
batch_4cpu	icl spp2	None	sub2
batch_8cpu	icl spp2	None	sub2
batch_k580	icl580	None	-----
cmc	icl spp2	None	sub2
interact_spp2	icl spp2	None	sub1
interact_k580	icl580	None	-----

La tabella 2 riporta le caratteristiche delle code batch che mandano in esecuzione job sulle macchine icl spp2 e icl580, accessibili a tutti gli utenti, in termini di:

- tipo di job da sottoporre alla coda (parallelo/sequenziale)
- durata del job (breve/nessun limite sul tempo)



- numero di job che possono essere in esecuzione contemporaneamente (MAX SLOT, slot si deve intendere uno per ogni cpu utilizzata, quindi 1 nel caso di job sequenziali, 4 per es. nel caso di job paralleli a 4 cpus)
- numero di job appartenenti allo stesso utente che possono essere in esecuzione contemporaneamente (JOB SLOT per USER)
- numero massimo di processori utilizzabili da ciascun job (PROCLIMIT)
- file system scratch da utilizzare (DIR)

Come sottomettere una richiesta ad una coda batch

Se, ad esempio, si vuole sottomettere una richiesta alla coda batch_seq si deve procedere nel seguente modo:

- la coda batch_seq è associata al subcomplex sub1 e quindi all'area di scratch /scratch1; poiché le prestazioni della macchina sono decisamente migliori se si lavora sulle aree di scratch direttamente connesse all'ipernodo su cui è definito il subcomplex utilizzato, è necessario creare una directory di lavoro sotto /scratch1
- i comandi da eseguire in batch devono essere contenuti in uno shell script, ad esempio lancia.sh di questo tipo:

```
.....file
lancia.sh.....
#!/bin/csh
cd /scratch1/<directory_di_lavoro>
<comandi_da_eseguire_in_batch>
.....
```

ATTENZIONE: il file lancia.sh deve essere eseguibile.

Per sottomettere la richiesta alla coda batch_seq si utilizza il comando:

bsub -q batch_seq lancia.sh

Si consiglia comunque di utilizzare la forma:

bsub -q batch_seq -o file.out -e file.err lancia.sh

dove file.out è l'eventuale file di output del batch e file.err è il file dove dirottare gli eventuali errori.

Come controllare lo stato delle code

Per controllare lo stato complessivo delle code si utilizza il comando: ***bqueues***

Per controllare lo stato dei propri job il comando è: ***bjobs***

Con l'opzione -p da' solo l'elenco dei job pending e il motivo per cui sono in pending.

Con l'opzione -s dà solo l'elenco dei job sospesi e il motivo per cui sono sospesi.

Se si aggiunge l'opzione -u all
dà l'elenco di tutti i jobs sulla macchina, non solo dei propri.

Per deletare un job il comando da utilizzare è:

bkill <job-id>

Per una completa descrizione dei comandi bsub, bqueues, bjobs, bkill e dei relativi parametri si rimanda al manuale in linea (man bsub, man bqueues, man bjobs, man bkill).

Si ricorda che la maggior parte dei programmi applicativi sono dotati di una propria sintassi per l'utilizzo in modalità batch che prevede già l'utilizzo del nuovo sistema LSF.

Si rimanda agli help in linea per le relative informazioni

(v. file /users/PACCHETTI.INSTALLATI).

Tabella 2

Coda	Tipo job	Durata job	MAX SLOT	JOB SLOT USER	PROCLIMIT	DIR
short1	Seq	15	2	1	1	/scratch1
Short2	Par	15	16	8	8	/scratch2
Short_k580	Seq/par	15	8	4	4	/scratch1
Batch_seq	Seq	No_limit	None	3	1	/scratch1
Batch_4cpu	Par	No_limit	16	4	4	/scratch2
Batch_8cpu	Par	No_limit	16	8	8	/scratch2
Batch_k580	Seq/par	No_limit	8	4	4	/scratch1